ENVIRONMETRICS
Environmetrics, **10**, 725–736 (1999)

CHANGE-POINTS AND BOOTSTRAP

EDIT GOMBAY[1]* AND LAJOS HORVÁTH[2]

[1]*Department of Mathematical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1*
[2]*Department of Mathematics, University of Utah Salt Lake City, UT 84112, USA*

SUMMARY

We show that the weighted bootstrap can be used to detect possible changes in the distribution of random vectors. We illustrate our method with change-point detection in the monthly precipitation and water discharges from Málo Ráztoka. Copyright © 1999 John Wiley & Sons, Ltd.

KEY WORDS water discharge; change-point; bootstrap; empirical distribution function; Kendall's tau

1. INTRODUCTION

Jarušková (1994) and Gombay and Horváth (1997) analyzed the monthly averages of water discharges from Načetínský measured during 1951–1990. Načetínský is located in the Erzgebirge mountains and it was expected that the large deforestation in the Erzgebirge mountains may have changed water discharges from Načetínský. Jarušková (1994) and Gombay and Horváth (1997) assumed that the data were from log-normal distribution. Changes in the mean of the log-transformed data were found. In this paper we are looking for possible changes in the monthly average water discharges from Málo Ráztoka and the monthly precipitation measured at the creek. No parametric form is assumed on the distributions. We discuss two methods to detect possible changes in the data. The first one is based on empirical distributions, while the second one uses Kendall's tau. We obtain limit theorems for the test statistics under the 'no change' null hypothesis. In both cases the limits will depend on the unknown underlying distribution, so the results cannot be applied immediately to analyze the Málo Ráztoka data. We use the weighted bootstrap to approximate the distributions of the test statistics.

2. EMPIRICAL DISTRIBUTIONS

Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be independent random variables with distribution functions $F^{(1)}(x,y), F^{(2)}(x,y), \ldots, F^{(n)}(x,y)$. We wish to test the 'no change' null-hypothesis

$$H_0 : F^{(1)}(x, y) = F^{(2)}(x, y) = \cdots = F^{(n)}(x, y) \text{ for all } (x, y) \in R^2$$

* Correspondence to: E. Gombay, Department of Mathematical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1.

Contract grant sponsor: NSERC Canada grant.
Contract grant sponsor: NATO Grant.
Contract grant number: CRG 960503.

CCC 1180–4009/99/060725–12$17·50
Received 1 December 1998
Accepted 13 April 1999

against the alternative

$$H_A: \text{ there is an integer } k^*, \quad 1 \leqslant k^* < n \text{ such that } F^{(1)}(x,y) = \cdots = F^{(k^*)}(x,y),$$
$$F^{(k^*+1)}(x,y) = \cdots = F^{(n)}(x,y) \text{ and } F^{(k^*)}(x_0,y_0) \neq F^{(k^*+1)}(x_0,y_0) \text{ with some } (x_0,y_0).$$

We follow the method of Csörgö and Horváth (1997, Section 2.6). We divide the random sample into two parts $\{(X_i, Y_i), 1 \leqslant i \leqslant k\}$, $\{(X_i, Y_i), k < i \leqslant n\}$ and compute the corresponding empirical distribution functions

$$\hat{F}_i(x,y) = \frac{1}{k} \sum_{1 \leqslant i \leqslant k} I\{X_i \leqslant x, \, Y_i \leqslant y\}$$

and

$$\hat{F}_k^*(x,y) = \frac{1}{n-k} \sum_{k \leqslant i \leqslant n} I\{X_i \leqslant x, \, Y_i \leqslant y\}.$$

We reject H_0 if

$$\hat{T}_n = \max_{1 \leqslant k < n} \frac{k(n-k)}{n^{3/2}} \sup_{x,y} | \hat{F}_k(x,y) - F_k^*(x,y) |$$

is large. Let F denote the common distribution function under H_0. The limit of $\hat{T}_n$ is given by

$$\xi = \sup_{0 \leqslant t \leqslant 1} \sup_{x,y} |\Gamma_F(x,y;t)|,$$

where $\{\Gamma_F(x,y;t), \; (x,y) \in R^2, \; t \in R\}$ is a Gaussian process with $E\Gamma_F(x,y;t) = 0$ and $E\Gamma_F(x,y;t)\Gamma_F(x',y';t') = \{(t \wedge t') - tt'\}\{F(x \wedge x', y \wedge y') - F(x,y)F(x',y')\}$ where $a \wedge b = \min(a,b)$. The next result was obtained by Csörgö and Horváth (1997, p. 153).

Theorem 2.1. *If H_0 holds, then*

$$\hat{T}_n \xrightarrow{D} \xi \quad (n \to \infty).$$

Csörgö and Horváth (1997) pointed out that the test based on $\hat{T}_n$ has higher power against change in the middle than against early or late changes. Introducing weight functions we can increase the power of the tests against early or late changes. For details we refer to Csörgö and Horváth (1997).

Unlike in the case of the univariate empirical process, the distribution of the supremum of the weak limit of the bivariate empirical process depends on the underlying distribution function even under the null-hypothesis. The underlying distribution is not specified under the null-hypothesis, so we cannot use Monte-Carlo simulations to approximate the distribution function of ξ. Thus in the applications Theorem 2.1 cannot be used directly to approximate the critical values of $\hat{T}_n$. However, we show that the weighted bootstrap can be used to approximate the distribution of $\hat{T}_n$. We assume that

$$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n \text{ are independent, identically distributed random variables with var } \varepsilon_1 = 1 \text{ and } E\varepsilon_1^4 < \infty \tag{1}$$

and

$$\{\varepsilon_i, \quad 1 \leqslant i \leqslant n\} \text{ and } \{(X_i, Y_i), \ 1 \leqslant i \leqslant n\} \text{ are independent.} \tag{2}$$

Let

$$U_n(x, y; k) = \left(1 - \frac{k}{n}\right) n^{-1/2} \sum_{1 \leqslant i \leqslant k} (\varepsilon_i - \hat{\varepsilon}_k) I\{X_i \leqslant x, Y_i \leqslant y\} - \frac{k}{n} n^{-1/2} \sum_{k < i \leqslant n} (\varepsilon_i - \hat{\varepsilon}_k^*) I\{X_i \leqslant x, Y_i \leqslant y\}, \tag{3}$$

$U_n(x, y; 0) = U_n(x, y, n) = 0$, where

$$\hat{\varepsilon}_i = \frac{1}{k} \sum_{1 \leqslant i \leqslant k} \varepsilon_i \text{ and } \hat{\varepsilon}_k^* = \frac{1}{n-k} \sum_{k < i \leqslant n} \varepsilon_i.$$

We use replicas of

$$\tilde{T}_n = \max_{1 \leqslant k < n} \sup_{x,y} |U_n(x, y; k)|$$

to approximate the distribution of $\hat{T}_n$. Our next result shows that $\hat{T}_n$ and $\tilde{T}_n$ have the same limit distribution.

Theorem 2.2. *If H_0, (1) and (2) hold, then*

$$\tilde{T}_n \to \xi \quad (n \to \infty).$$

The proof of Theorem 2.2 will be given in Section 5.

Now the simulation of the critical values of $\hat{T}_n$ can be implemented. We generate N independent copies of the εs, $\{\varepsilon_{i,1}, 1 \leqslant i \leqslant n\}, \ldots, \{\varepsilon_{i,N}, 1 \leqslant i \leqslant n\}$ and using $\{\varepsilon_{i,j}, 1 \leqslant i \leqslant n\}$ we compute $T_{n,j,}$ $1 \leqslant j \leqslant N$. Let

$$H_{N,n}(t) = \frac{1}{N} \sum_{1 \leqslant i \leqslant N} I\{\tilde{T}_{n,j} \leqslant t\}.$$

Putting together Theorems 2.1 and 2.2 we obtain that under the null-hypothesis

$$\sup_t |P\{\hat{T}_n \leqslant t\} - H_{N,n}(t)| \xrightarrow{P} 0 \quad (N \wedge n \to \infty). \tag{4}$$

Thus we have established that the data-driven simulation gives a uniformly consistent estimator for the distribution function of $\hat{T}_n$ under H_0. In order to have reasonable power when a change occurred in the data we must consider the behaviour of $\hat{T}_n$ and $\tilde{T}_n$ under the alternative. It turns our that $\hat{T}_n$ and $\tilde{T}_n$ have different order under H_A. The consistency of our procedure will be an immediate consequence of the following result.

Theorem 2.3. *If H_A, (1) (2) hold and $k^* = [n\theta]$ with some $0 < \theta < 1$, then*

$$\hat{T}_n \xrightarrow{P} \infty \quad (n \to \infty) \tag{5}$$

and

$$\tilde{T}_n = O_P(1) \quad (n \to \infty). \tag{6}$$

We prove Theorem 2.3 in Section 5.

Putting together Theorems 2.1–2.3 we conclude that the weighted bootstrap gives a rejection region which has the correct asymptotic significance level under the null-hypothesis and very high probability of rejection under the alternative. To demonstrate our claim, for any $0 < \alpha < 1$ we define $z_{N,n}(\alpha)$ by

$$z_{N,n}(\alpha) = \inf\{t : \; H_{N,n}(t) \geqslant 1 - \alpha\}.$$

Then by (4) we have under H_0 that

$$\limsup_{N \wedge n \to \infty} P\{\hat{T}_n \geqslant z_{N,n}(\alpha)\} \leqslant \alpha$$

while under H_A by Theorem 2.3 we conclude

$$\lim_{N \wedge n \to \infty} P\{\hat{T}_n \geqslant z_{N,n}(\alpha)\} = 1.$$

(We note that if F is continuous, then ξ has a continuous distribution function and therefore we have that $P\{\hat{T}_n \geqslant z_{N,n}(\alpha)\} \to \alpha$ as $N \wedge n \to \infty$.)

3. KENDALL'S TAU

Kendall's tau is a popular measure of association between random variables. Similarly to the first section we split the data into two sub-samples after X_k and define

$$\hat{\tau}_k = \frac{2}{k(k-1)} \sum_{1 \leqslant i < j \leqslant k} I\{(X_i - X_j)(Y_i - Y_j) > 0\}$$

and

$$\hat{\tau}_k^* = \frac{2}{(n-k)(n-k-1)} \sum_{k < i < j \leqslant n} I\{(X_i - X_j)(Y_i - Y_j) > 0\}.$$

We reject H_0, if

$$\hat{M}_n = \max_{1 \leqslant k < n} \frac{k(n-k)}{n^{3/2}} |\hat{\tau}_k - \hat{\tau}_k^*| \tag{7}$$

is large. We note that $\hat{M}_n$ is a maximally selected weighted difference between two bivariate U-statistics. For some applications of U-statistics to change-point analysis we refer to Csörgö and Horváth (1988) and Gombay and Horváth (1995) (cf. also Section 2.4 in Csörgö and Horváth 1997). Our first result is the limit distribution of $\hat{M}_n$ and H_0.

Theorem 3.1. *If H_0 holds, then*

$$\hat{M}_n \xrightarrow{D} \sigma \sup_{0 \leqslant t \leqslant 1} |B(t)| \quad (n \to \infty)$$

with some $\sigma = \sigma(F)$, where $\{B(t), 0 \leqslant t \leqslant 1\}$ stands for a Brownian bridge.

Since $\sigma = \sigma(F)$ depends on the unknown F it is difficult to estimate it from the sample. We use again the weighted bootstrap to approximate the distribution of $\hat{M}_n$. Let

$$\tilde{U}_n(k) = 2n^{-1/2}\left\{\frac{n-k}{kn} \sum_{1 \leqslant i,j \leqslant k} (\varepsilon_i - \hat{\varepsilon}_k) I\{(X_i - X_j)(Y_i - Y_j) > 0\} \right.$$
$$\left. - \frac{k}{n(n-k)} \sum_{k < i,j \leqslant n} (\varepsilon_i - \hat{\varepsilon}_k^*) I\{(X_i - X_j)(Y_i - Y_j) > 0\}\right\}$$

and

$$\tilde{M}_n = \max_{1 \leqslant k < n} |\tilde{U}_n(k)|.$$

Theorem 3.2. *If H_0 holds, then*

$$\tilde{M}_n \xrightarrow{D} \sigma \sup_{0 \leqslant t \leqslant 1} |B(t)| \quad (n \to \infty),$$

where $\sigma = \sigma(F)$ is from Theorem 3.1 and $\{B(t), 0 \leqslant t \leqslant 1\}$ stands for a Brownian bridge.

Putting together Theorems 3.1 and 3.2 we see that the weighted bootstrap gives a consistent estimator for the distribution function of $\hat{M}_n$ under H_0 when the scheme discussed in Section 2 is used. The consistency of the test procedure can be discussed similarly to Theorem 2.3. We note that using Kendall's tau we are checking if the parameter $P\{(X_i - X_j)(Y_i - Y_j) > 0\}$ is the same for the first k and the last $n - k$ observations for each k, $1 \leqslant k < n$.

4. APPLICATIONS AND SIMULATIONS

The Málo Ráztoka data consist of 36 pairs of monthly averages of water discharges and precipitations for each months. We considered the measurements for October. Figure 1 is the graph of $V_n(k) = n^{-3/2}k(n-k)\sup_{x,y} |\hat{F}_k(x, y) - \hat{F}_k^*(x, y)|$ for this data set. We used standard normal random variables to generate the bootstrapped statistics. The bootstrap procedure was repeated $N = 1000$ times. One of the graphs of the simulated $\sup_{x,y} | U_n(x, y; k)|$ is given in Figure 2. No significant change in the data was found at 1 per cent significance level.

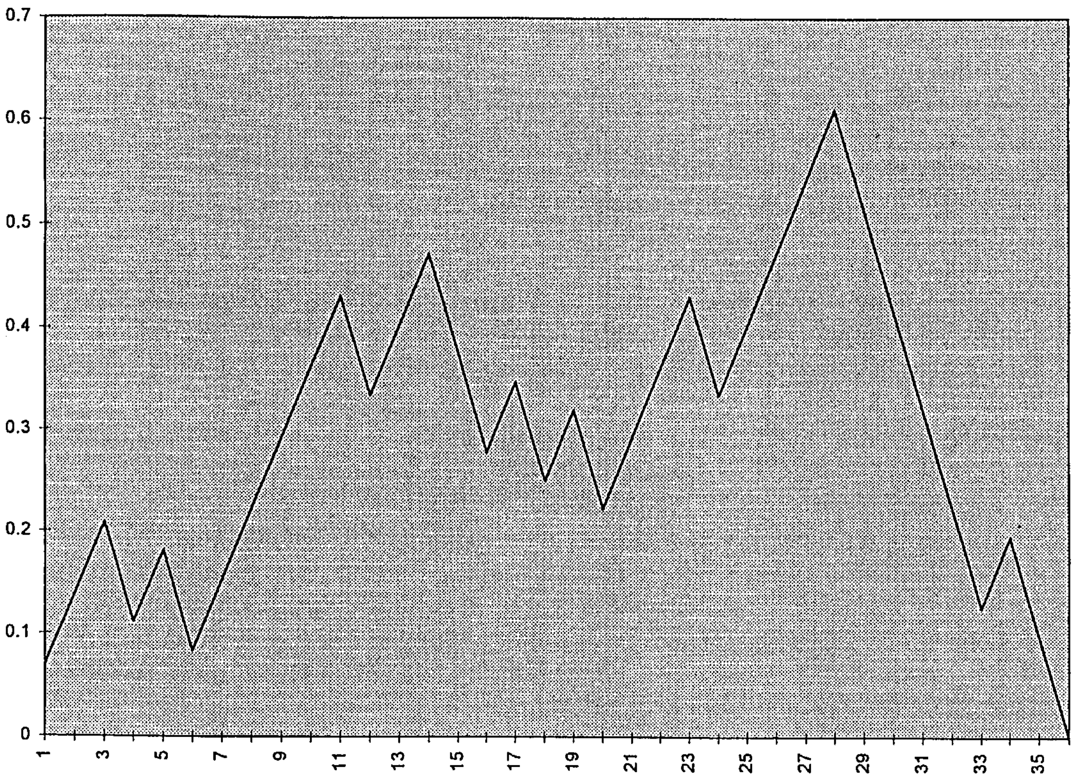

Figure 1. The graph of $V_n(k)$

We also computed $\hat{\tau}_k$ and $\hat{\tau}_k^*$ for the Málo Ráztoka data. Figure 3 is the graph of $R_n(k) = n^{-3/2}k(n-k)\,|\,\hat{\tau}_k - \hat{\tau}_k^*\,|$ for October. We used again standard normal weights in the bootstrap and the simulations were repeated $N = 1000$ times. A typical graph of $\tilde{U}_n(k)$ is given in Figure 4. No significant change was found at 1 per cent significance level.

We repeated our procedures for all months. No significant changes were found at 1 per cent significance levels.

A small simulation study was performed to compare the approximations for the distribution function of $\hat{M}_n$ provided by Theorem 3.1 and the bootstrap method in Theorem 3.2. If H_0 holds and X_1 and Y_1 are independent, then $\sigma = 1/3$ in Theorems 3.1 and 3.2. We generated $n = 50$ independent pairs of independent standard normal random variables. Using again standard normal weights and $N = 1000$ replicas the bootstrap approximation was computed for $P\{\hat{M}_{50} \leqslant t\}$. Figure 5 has the bootstrap approximation and $P\{\sup_{0\leqslant x\leqslant 1}|B(x)| \leqslant 3t\}$.

5. PROOFS

We recall that $F(x, y)$ denotes the common distribution under H_0. We can assume without loss of generality (cf. Wichura 1973) that $F(x,y)$ has uniform marginals on [0, 1]. Let $\mu = E\varepsilon_i$ and

$$Z_n(x, y; t) = n^{-1/2} \sum_{1 \leqslant i \leqslant nt} (\varepsilon_i - \mu) I\{X_i \leqslant x,\, Y_i \leqslant y\}.$$

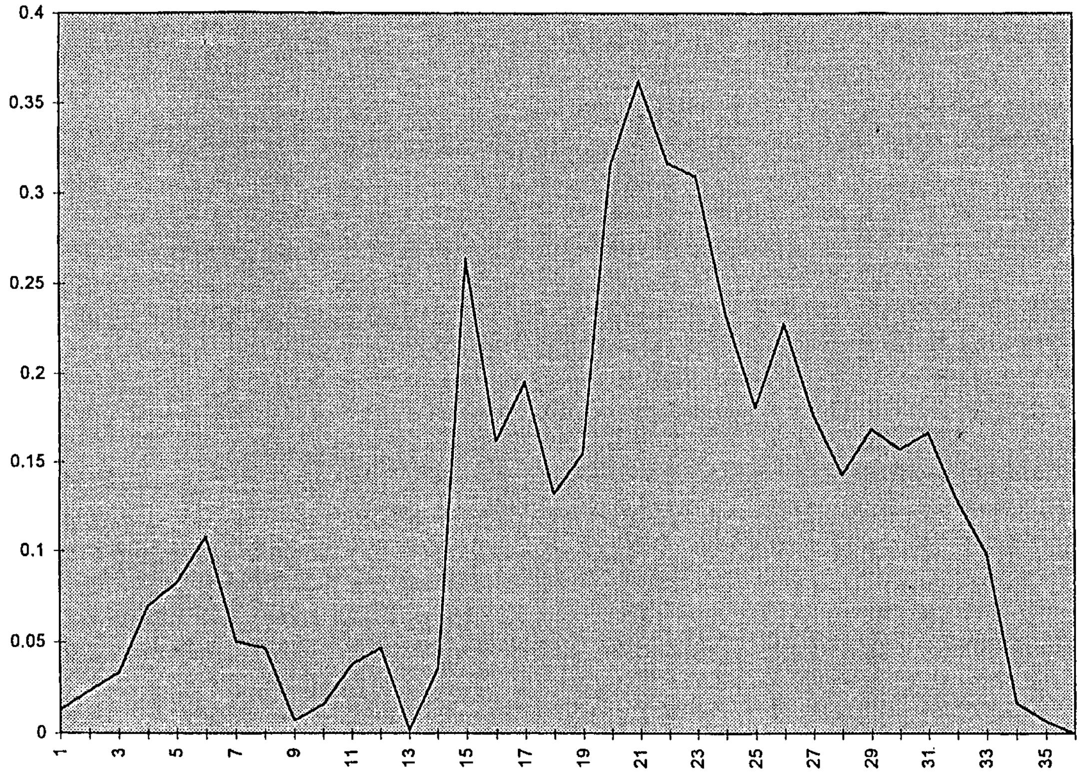

Figure 2. The graph of $\sup_{x,y} | U_n(x, y; k) |$ using standard normal weights

Lemma 5.1: If H_0, (1) and (2) hold, then

$$Z_n(x, y; t) \overset{\mathcal{D}[0,1]^3}{\rightarrow} W_F(x, y; t),$$

where $\{W_F(x, y; t), 0 \leqslant x, y, t \leqslant 1\}$ is a Gaussian process with $EW_F(x, y; t) = 0$ and $EW_F(x, y; t) W_F(x', y'; t') = (t \wedge t') F(x \wedge x', y \wedge y')$.

Proof: We apply Theorem 6 of Bickel and Wichura (1971). Let C and D be neighboring blocks in the unit square in the sense of Bickel and Wichura (1971). They showed, that there is a finite, continuous measure ν on the unit square such that

$$E(I\{(X_i, Y_i) \in C\})^2 \leqslant \nu(C)$$

and

$$E(I\{X_i, Y_i) \in C\})^2 (I\{X_i, Y_i) \in D\})^2 \leqslant 3\nu(C)\nu(D).$$

Hence using (1) and (2) we obtain that

$$E((\varepsilon_i - \mu) I\{(X_i, Y_i) \in C\})^2 = E(\varepsilon_i - \mu)^2 E(I\{X_i, Y_i) \in C\})^2 \leqslant \nu(C)$$

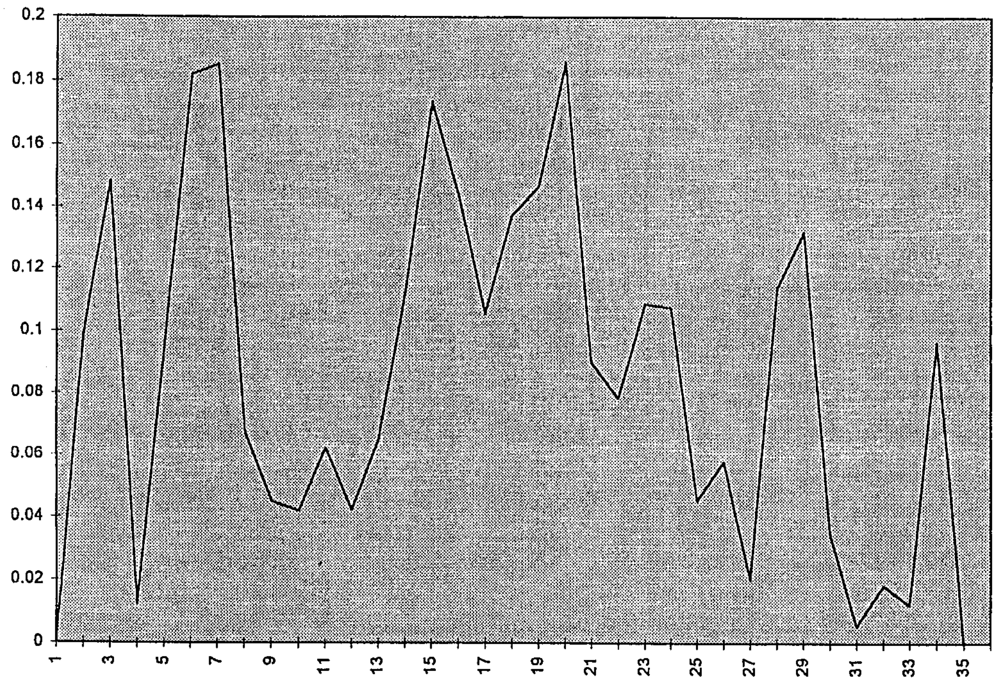

Figure 3. The graph of $R_n(k)$

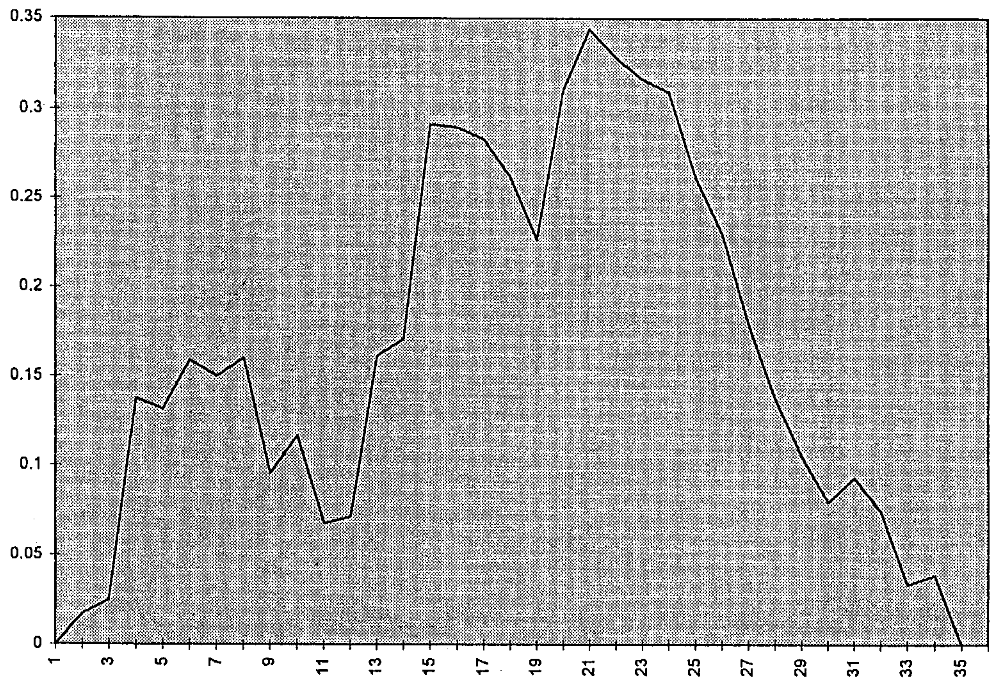

Figure 4. The graph of $\sup_{x,y} | \tilde{U}_n(x, y; k) |$ using standard normal weights

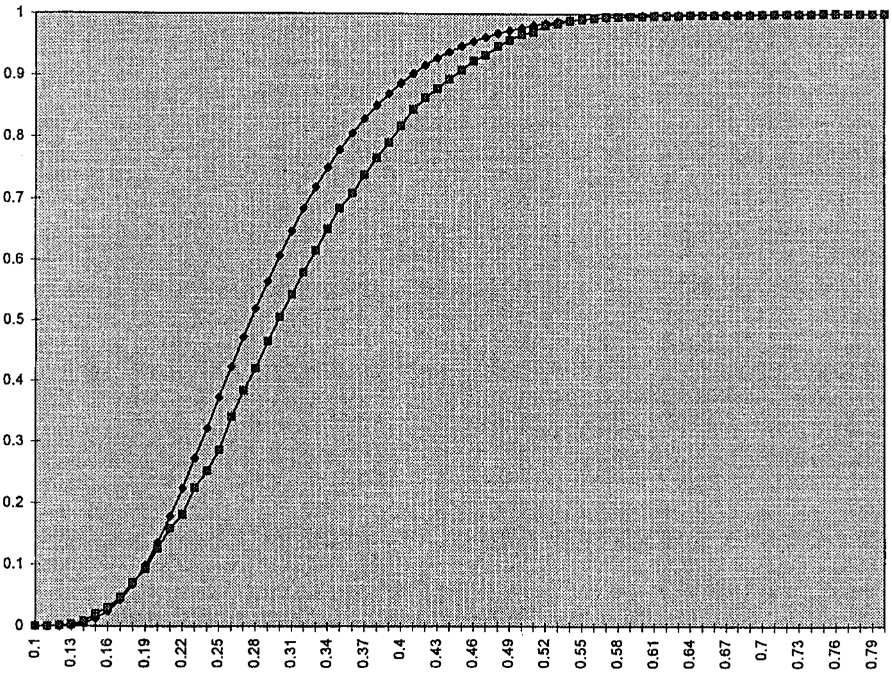

Figure 5. The graphs of the bootstrap approximation for $P\{\hat{M}_{50} \leqslant t\}$ (-■-), and $P\{\sup[_{0 \leqslant x \leqslant 1} |B(x)| \leqslant 3t\}$ (-◆-)

and

$$
\begin{aligned}
E((\varepsilon_i - \mu)I\{(X_i, Y_i) \in C\})^2((\varepsilon_i - \mu)I\{(X_i, Y_i) \in D\})^2 \\
&\leqslant E(\varepsilon_i - \mu)^4 E(I\{(X_i, Y_i) \in C\})^2(I\{(X_i, Y_i) \in D\})^2 \\
&\leqslant 3E(\varepsilon_1 - \mu)^4 \nu(C)\nu(D).
\end{aligned}
$$

Lemma 5.1 follows immediately from Theorem 6 of Bickel and Wichura (1971).

Proof of theorem 2.2: The process $U_n(x, y; k)$ does not depend on μ so we can assume that $\mu = 0$. First we write

$$
\begin{aligned}
U_n(x, y; k) = {} & \left(1 - \frac{k}{n}\right) \sum_{1 \leqslant i \leqslant k} \varepsilon_i I\{X_i \leqslant x, Y_i \leqslant y\} - \left(1 - \frac{k}{n}\right) F(x, y) n^{-1/2} \sum_{1 \leqslant i \leqslant k} \varepsilon_i \\
& - \frac{k}{n} n^{-1/2} \sum_{k < i \leqslant n} \varepsilon_i I\{X_i \leqslant x, Y_i \leqslant y\} + \frac{k}{n} F(x, y) n^{-1/2} \sum_{k < i \leqslant n} \varepsilon_i \\
& + \left(1 - \frac{k}{n}\right)\left(F(x, y) - \frac{1}{k} \sum_{1 \leqslant j \leqslant k} I\{X_j \leqslant x, Y_j \leqslant y\}\right) n^{-1/2} \sum_{1 \leqslant i \leqslant k} \varepsilon_i \\
& - \frac{k}{n}\left(F(x, y) - \frac{1}{n-k} \sum_{k < j \leqslant n} I\{X_j \leqslant x, Y_j \leqslant y\}\right) n^{-1/2} \sum_{k < i \leqslant n} \varepsilon_i. \qquad (8)
\end{aligned}
$$

By the law of the interated logarithm we have

$$\max_{1 \leqslant k < n} \sup_{0 \leqslant x,y \leqslant 1} \left| \left(1 - \frac{k}{n}\right) \left(F(x, y) - \frac{1}{k} \sum_{1 \leqslant j \leqslant k} I\{X_j \leqslant x,\ Y_j \leqslant y\}\right) n^{-1/2} \sum_{1 \leqslant i \leqslant k} \varepsilon_i \right|$$
$$= O_P(n^{-1/2} \log \log n) \tag{9}$$

and

$$\max_{1 \leqslant k < n} \sup_{0 \leqslant x,y \leqslant 1} \left| \frac{k}{n} \left(F(x, y) - \frac{1}{n-k} \sum_{k < j \leqslant n} I\{X_j \leqslant x,\ Y_j \leqslant y\}\right) n^{-1/2} \sum_{k < i \leqslant n} \varepsilon_i \right|$$
$$= O_P(n^{-1/2} \log \log n). \tag{10}$$

Hence Lemma 5.1 yields

$$U_n(x, y; nt) \overset{D[0,1]^3}{\longrightarrow} \Gamma_F(x, y; t),$$

where

$$\Gamma_F(x, y; t) = (1 - t)\{W_F(x, y; t) - F(x, y) W_F(1, 1; t)\} - t\{W_F(x, y; 1) - W_F(x, y; t)$$
$$- F(x, y)(W_F(1, 1; 1) - W_F(1, 1; t))\}. \tag{c}$$

It is easy to see that $\Gamma_F(x, y; t)$ is Gaussian with $E\Gamma_F(x, y; t) = 0$ and $E\Gamma_F(x, y; t)\Gamma_F(x', y'; t') = (t \wedge t' - tt')(F(x \wedge x', y \wedge y') - F(x, y)F(x', y'))$.

Proof of theorem 2.3: Csörgö and Horváth (1997, Section 2.6) proved (5).

Since (8)–(10) hold under the alternative, it is enough to show that

$$\sup_{0 \leqslant x,y,t \leqslant 1} |Z_n(x, y; t)| = O_P(1). \tag{11}$$

First we note that

$$\sup_{0 \leqslant x,y,t \leqslant 1} |Z_n(x, y; t)| \leqslant \sup_{1 \leqslant k \leqslant k^*} \sup_{x,y} n^{-1/2} \left| \sum_{1 \leqslant i \leqslant k} (\varepsilon_i - \mu) I\{X_i \leqslant x,\ Y_i \leqslant y\} \right|$$
$$+ \sup_{k^* \leqslant k \leqslant n} \sup_{x,y} n^{-1/2} \left| \sum_{k < i \leqslant n} (\varepsilon_i - \mu) I\{X_i \leqslant x,\ Y_i \leqslant y\} \right|.$$

Following the proof of Lemma 5.1 one can establish that $\{n^{-1/2}\Sigma_{1 \leqslant i \leqslant nt})(\varepsilon_i - \mu) I\{X_i \leqslant x, Y_i \leqslant y\}$, $0 \leqslant t \leqslant \theta,\ 0 \leqslant x,\ y \leqslant 1\}$ and $\{n^{-1/2}\Sigma_{nt \leqslant i \leqslant n}(\varepsilon_i - \mu) I\{X_i \leqslant x,\ y_i \leqslant y\},\ \theta \leqslant t \leqslant 1,\ 0 \leqslant x,\ y \leqslant 1\}$ converge weakly. Hence (11) is proven.

Proof of Theorem 3.1: The result is a bivariate version of Theorem 1.1 of Gombay and Horváth (1995) [cf. also Theorem 2.4.7 in Csörgö and Horváth (1997)]. Using Hall (1979) we get that

$$\max_{1 \leqslant k < n} \left| \frac{k(n-k)}{n^{3/2}} (\hat{\tau}_k - \hat{\tau}_k^*) - 2 \int\int I\{(x_1 - x_2)(y_1 - y_2) > 0\} \right. \\ \left. \mathrm{d}F(x_1, y_1)\, \mathrm{d}\left\{ \frac{k(n-k)}{n^{3/2}} (\hat{F}_k(x_2, y_2) - F_k^*(x_2, y_2)) \right\} \right| = O_P(1), \tag{12}$$

and therefore without any modification one can copy the proof of Theorem 1.1 of Gombay and Horváth (1995). The details are omitted.

Proof of Theorem 3.2: The proof is rather technical and lengthy so we will sketch the major steps only. First we note that

$$\tilde{U}_n(k) = 2n^{-3/2}(n-k) \int\int I\{(x_1 - x_2)(y_1 - y_2) > 0\} \mathrm{d}\hat{F}_i(x_1, y_1) \\ \mathrm{d}\left\{ \sum_{1 \leqslant i \leqslant k} (\varepsilon_i - \hat{\varepsilon}_k) I\{X_i \leqslant x_2,\, Y_1 \leqslant y_2\} \right\} - 2n^{-3/2} k \int\int I\{(x_1 - x_2)(y_1 - y_2) > 0\} \\ \mathrm{d}F_k^*(x_1, y_1) \mathrm{d}\left\{ \sum_{k < i \leqslant n} (\varepsilon_i - \hat{\varepsilon}_k^*) I\{X_i \leqslant x_2,\, Y_i \leqslant y_2\} \right\}.$$

Next we must show that $\hat{F}_k(x_1, y_1)$ and $F_k^*(x_1, y_1)$ can be replaced with $F(x_1, y_1)$ so we could conclude

$$\max_{1 \leqslant k < n} \left| \tilde{U}_n(k) - 2 \int\int I\{(x_1 - x_2)(y_1 - y_2) > 0\} \mathrm{d}F\,(x_1, y_1) \right. \\ \mathrm{d}\left\{ n^{-1/2}\left(1 - \frac{k}{n}\right) \sum_{1 \leqslant i \leqslant k} (\varepsilon_i - \hat{\varepsilon}_k) I\{X_i \leqslant x_2,\, Y_i \leqslant y_2\} \right. \\ \left. \left. - n^{-1/2} \frac{k}{n} \sum_{k < i \leqslant n} (\varepsilon_i - \hat{\varepsilon}_i) I\{X_i \leqslant x_2,\, Y_i \leqslant y_2\} \right\} \right| = O_p(1). \tag{13}$$

We showed in the proof of Theorem 2.1 that $n^{-3/2}[nt](n - [nt])(\hat{F}_{[nt]}(x, y) - F^*_{[nt]}(x, y))$ and $n^{-3/2}$ $\{(n - [nt])\Sigma_{1 \leqslant i \leqslant [nt]}(\varepsilon_i - \hat{\varepsilon}_{[nt]}) I\{X_i \leqslant x, Y_i \leqslant y\} - [nt]\Sigma_{[nt] < i \leqslant n}(\varepsilon_i - \hat{\varepsilon}^*_{[nt]}) I\{X_i \leqslant x, Y_i \leqslant y\}\}$ converge weakly to the same process, so in light of (12) and (13), it is heuristically clear that the limits in Theorems 3.1 and 3.2 must be the same.

REFERENCES

Bickel, P. J. and Wichura, M. J. (1971). 'Convergence criteria for multi-parameter stochastic processes and some applications'. *Annals of Mathematics and Statistics* **42**, 1656–1670.

Csörgö, M. and Horváth, L. (1988). 'Invariance principles for changepoint problems'. *Journal of Multivariate Analysis* **27**, 151–168.

Csörgö, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*. Chichester: Wiley.

Gombay, E. and Horváth, L. (1995). 'An application of *U*-statistics to change-point analysis'. *Acta Scientiarum Mathematicarum (Szeged)* **60**, 345–357.

Gombay, E. and Horváth, L. (1997). 'An application of the likelihood method to change-point detection'. *Environmetrics* **8**, 459–467.

Hall, P. (1979). 'On the invariance principle for *U*-statistics'. *Stochastic Processes and their Applications* **9**, 163–174.

Jarušková, D. (1994). 'Applications of change-point detection to ecology'. Preprint.

Wichura, M. J. (1973). 'Some Strassen-type laws of the iterated logarithm for multi-parameter stochastic processes with independent increments'. *Annals of Probability* **1**, 272–296.

